BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 4: e59082
doi: 10.3897/biss.4.59082

OPEN ACCESS

Conference Abstract

Change in Pictures: Creating best practices in archiving ecological imagery for reuse

Corinna Gries[‡], Stace Beaulieu[§], Renée F. Brown[|], Gastil Gastil-Buhl[¶], Sarah Elmendorf[#], Hsun-Yi Hsieh[¤], Li Kui[¶], Greg Maurer[«], John H. Porter[»]

‡ University of Wisconsin, Madison, Madison, United States of America
§ Woods Hole Oceanographic Institution, Woods Hole, United States of America
| University of New Mexico, Albuquerque, United States of America
¶ University of California, Santa Barbara, Santa Barbara, United States of America
# University of Colorado, Boulder, United States of America
¤ Michigan State University, Hickory Corners, United States of America
« New Mexico State University, Las Cruces, United States of America
» University of Virginia, Charlottesville, United States of America

Corresponding author: Corinna Gries (cgries@wisc.edu)

Received: 28 Sep 2020 | Published: 30 Sep 2020

Citation: Gries C, Beaulieu S, Brown RF, Gastil-Buhl G, Elmendorf S, Hsieh H-Y, Kui L, Maurer G, Porter JH (2020) Change in Pictures: Creating best practices in archiving ecological imagery for reuse. Biodiversity Information Science and Standards 4: e59082. https://doi.org/10.3897/biss.4.59082

Abstract

The research data repository of the Environmental Data Initiative (EDI) is building on over 30 years of data curation research and experience in the National Science Foundation-funded US Long-Term Ecological Research (LTER) Network. It provides mature functionalities, well established workflows, and now publishes all 'long-tail' environmental data. High quality scientific metadata are enforced through automatic checks against community developed rules and the Ecological Metadata Language (EML) standard. Although the EDI repository is far along in making its data findable, accessible, interoperable, and reusable (FAIR), representatives from EDI and the LTER are developing best practices for the edge cases in environmental data publishing. One of these is the vast amount of imagery taken in the context of ecological research, ranging from wildlife camera traps to plankton imaging systems to aerial photography. Many images are used in biodiversity research for community analyses (e.g., individual counts, species cover, biovolume, productivity), while others are taken to study animal behavior and landscape-level change.

© Gries C et al. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Some examples from the LTER Network include: using photos of a heron colony to measure provisioning rates for chicks (Clarkson and Erwin 2018) or identifying changes in plant cover and functional type through time (Peters et al. 2020). Multi-spectral images are employed to identify prairie species. Underwater photo quads are used to monitor changes in benthic biodiversity (Edmunds 2015). Sosik et al. (2020) used a continuous Imaging FlowCytobot to identify and measure phyto- and microzooplankton. Cameras at McMurdo Dry Valleys assess snow and ice cover on Antarctic lakes allowing estimation of primary production (Myers 2019).

It has been standard practice to publish numerical data extracted from images in EDI; however, the supporting imagery generally has not been made publicly available. Our goal in developing best practices for documenting and archiving these images is for them to be discovered and re-used. Our examples demonstrate several issues. The research questions, and hence, the image subjects are variable. Images frequently come in logical sets of time series. The size of such sets can be large and only some images may be contributed to a dedicated specialized repository. Finally, these images are taken in a larger monitoring context where many other environmental data are collected at the same time and location.

Currently, a typical approach to publishing image data in EDI are packages containing compressed (ZIP or tar) files with the images, a directory manifest with additional image-specific metadata, and a package-level EML metadata file. Images in the compressed archive may be organized within directories with filenames corresponding to treatments, locations, time periods, individuals, or other grouping attributes. Additionally, the directory manifest table has columns for each attribute. Package-level metadata include standard coverage elements (e.g., date, time, location) and sampling methods. This approach of archiving logical ‘sets’ of images reduces the effort of providing metadata for each image when most information would be repeated, but at the expense of not making every image individually searchable. The latter may be overcome if the provided manifest contains standard metadata that would allow searching and automatic integration with other images.

Keywords

data repository, metadata, ecological data

Presenting author

Corinna Gries

Presented at

TDWG 2020

Funding program

LTER

Author contributions

Authors have all contributed equally to writing this abstract.

References

- Clarkson C, Erwin R (2018) A time series of images of egret and cormorant colonies on Chimney Pole Marsh, VA 2010-2011 ver 8. Environmental Data Initiative. URL: https://doi.org/10.6073/pasta/cbd7550a71fc349d760a3ead334d4ec7
- Edmunds P (2015) MCR LTER: Coral Reef: Computer Vision: Multi-annotator Comparison of Coral Photo Quadrat Analysis. 3. Environmental Data Initiative. URL: https://doi.org/10.6073/pasta/8e09736f3c37c3554af349f21184b2a8
- Myers ME (2019) Spatiotemporal impact of snow on underwater photosynthetically active radiation in Taylor Valley, East Antarctica. LSU URL: https://digitalcommons.lsu.edu/gradschool_theses/4965
- Peters DC, Okin G, Herrick J, Savoy H, Anderson J, Scroggs SP, Zhang J (2020) Modifying connectivity to promote state change reversal: the importance of geomorphic context and plant–soil feedbacks. Ecology https://doi.org/10.1002/ecy.3069
- Sosik HM, Peacock E, Santos M (2020) Abundance and biovolume of taxonomically-resolved phytoplankton and microzooplankton imaged continuously underway with an Imaging FlowCytobot along the NES-LTER Transect in winter 2018. 1. Environmental Data Initiativ. URL: https://doi.org/10.6073/pasta/74775c4af51c237f2a20e4a8c011bc53